# MODELING MOBILE ADVERTISING CLICK FRAUD PROBLEM USING TENSORFLOW

Guoying (Grace) Zhang, Midwestern State University, 3410 Taft Blvd., Wichita Falls, TX 76310, (940)397-3289, grace.zhang@mwsu.edu

## RESEARCH ABSTRACT

This study intends to detect mobile advertisement click fraud using the tool box provided by Google's TensorFlow. Otentaryo et al. (2014) extended a research call on analyzing click fraud data from a real world mobile advertising network. The main task for their study is to identify fraudulent mobile application publishers. After careful investigation on the problem, it is believed that fraud identification should also be conducted at the click level instead of publisher level. Our goal is to effectively detect fraudulent clicks. The mobile network publisher in this study is BuzzCity Pte. Ltd., a global mobile advertising network that has millions of consumers around the world. The dataset we need to collect is a subset of the dataset from BuzzCity's sample quarterly data: "over 45 billion ad banners were delivered across the BuzzCity network, having over 10,000 publisher sites and reaching an average of 300 million unique users per month." (Otentaryo et al. 2014). In this particular study, we will train our model based a daily ad click stream set, and use another two daily click stream set for validation and testing. The entire dataset contains about 10 million mobile advertisement click information including: public IP address, device type, publisher information, ad campaign information, user country location, timestamp of this given click, and referred url address if exists.

TensorFlow is an open source software library for numerical computation using data flow graphs (TensorFlow, 2017). It is originally developed by Google's Brain Team for conducting machine learning and deep neural networks research. Since this is an open source system and has been updated continually in the developer's community, lots of domain applications, such as "speech recognition, computer vision, robotics, information retrieval, natural language processing, geographic information extraction, and computational drug discovery," and so on have chosen to use TensorFlow system to conduct research (Abadi et al. 2015)

## REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al. (2015) TensorFlow: Large-scale machine learning on heterogeneous systems. http://download.tensorflow.org/paper/whitepaper2015.pdf.

Otentaryo et al. (2014) Detecting Click Fraud in Online Advertising: A Data Mining Approach, *Journal of Machine Learning Research*, 15, 99-140

TensorFlow. (2017) About TensorFlow. https://www.tensorflow.org/ .